



XS-2130: MODELOS DE REGRESIÓN APLICADOS – GRUPO 01

PROGRAMA

II SEMESTRE 2018

Docente: Ricardo Alvarado Barrantes

Correo electrónico: estad.ucr@gmail.com

Oficina: 17 Estadística.

Teléfono: 25116529 / 84021263

Horario de clases: L: 9:00-11:50
J: 9:00-10:50

Horas de consulta: L: 3:00-5:00 pm
K: 1:00-3:00 pm

1. Descripción

Los modelos de regresión son un tema central en la práctica de la estadística y forman la base de un amplio rango de métodos estadísticos. En este curso se presentarán las bases de los modelos de regresión múltiple y los procedimientos desarrollados para asegurar una aplicación correcta de los mismos.

Además de los conocimientos teóricos se brindará al estudiante la posibilidad de aplicar los métodos mediante el uso del lenguaje de programación R.

- **Requisitos:** XS-2110 Métodos Estadísticos, XS-2210 Estadística Computacional I y MA-0313.
- **Correquisitos:** XS-2330 Estadística Computacional II.
- **Horas:** 5 horas semanales (2 de teoría y 3 de práctica).
- **Créditos:** 4.

2. Objetivo General

Proveer a los estudiantes el conocimiento conceptual y los instrumentos estadísticos para la aplicación de los modelos de regresión lineal.

3. Objetivos Específicos

Al finalizar el curso el estudiante tendrá criterio y conocimiento básico para:

- Reconocer las situaciones donde se pueden aplicar las técnicas de regresión lineal.
- Utilizar un modelo con fines de predicción (un valor estimado individual, un valor estimado promedio).
- Usar diferentes técnicas de selección de variables para formar un modelo apropiado a partir de un conjunto de predictores disponibles.
- Llevar a cabo los diagnósticos del modelo de regresión ajustado.
- Aplicar las medidas apropiadas en los casos de no conformidad con los supuestos del modelo de regresión.
- Analizar casos con respuestas de dos categorías y conteos.



4. Contenidos

I. El modelo de regresión lineal.	
1.1	El modelo y los supuestos.
1.2	Estimación de los coeficientes (notación matricial).
1.3	Valores ajustados y residuales.
1.4	Coefficientes estandarizados.
1.5	Inferencias. El cuadrado de medio de error. Varianzas de los estimadores. Intervalos de confianza para los coeficientes.
1.6	Intervalos de confianza para la respuesta media y predicción de nuevas observaciones.
1.7	Coefficientes de determinación (múltiple y parcial).
1.8	Comparación de modelos anidados con la distribución F. Uso de la distribución t.
1.9	Predictores cualitativos. Modelos con interacciones.
II. Selección de predictores.	
2.1	Procedimiento con todos los posibles subconjuntos: R^2 ajustado, suma de cuadrados de predicción (PRESS) y estadístico de Mallows (C_p).
2.2	Selección por pasos: criterios de información (Akaike y Bayes).
III. Diagnósticos.	
3.1	Verificación de los supuestos: a. No autocorrelación: prueba de Durbin-Watson. b. Normalidad: QQ plot. c. Homoscedasticidad: prueba de Breush Pagan. d. No multicolinealidad: factor de inflación de la variancia (VIF). e. Linealidad entre predictor y respuesta: gráficos de regresión parcial.
3.2	Medidas para detectar valores extremos: residuales estudentizados, "leverage".
3.3	Medidas para detectar casos de influencia: DFFITs, distancia de Cook, DFBETAs.
IV. Construcción del modelo.	
4.1	Transformaciones de Box-Cox.
4.2	Mínimos cuadrados ponderados.
4.3	Regresión robusta.
4.4	Componentes principales.
4.5	Validación del modelo.
V. Modelos lineales generalizados.	
5.1	Regresión logística.



5. Metodología

- Presentaciones teóricas: se impartirán lecciones magistrales por parte del docente donde se explicarán los conceptos y sus aplicaciones.
- Prácticas: se usará el laboratorio de computadoras de la escuela de estadística para el desarrollo de lecciones prácticas durante el semestre. Durante las sesiones de laboratorio se utilizará el lenguaje de programación R para realizar ejercicios de la materia vista en clase. Se realizarán ejercicios y sus respectivas prácticas.
- Tareas: se asignarán tareas y ejercicios para asegurar la “puesta en práctica” de los conceptos estudiados. Las tareas incluirán aplicaciones con datos para ser analizados con el programa elegido así como interpretaciones de los resultados
- Trabajo con datos reales:
 - Los estudiantes realizarán un trabajo de análisis de datos reales utilizando lo aprendido en el curso.
 - El trabajo será en grupos de 3 (máximo).
 - Las aplicaciones podrán ser tomadas de investigaciones en áreas como la economía, medicina, psicología, biología, población, etc.
 - Los estudiantes presentarán un anteproyecto donde expondrán los objetivos del estudio, el cual será criticado por el profesor y devuelto para que se realicen las mejoras sugeridas.
 - Al final de curso se espera que los estudiantes escriban un documento donde se describa detalladamente el procedimiento de análisis de los datos seleccionados.
 - El trabajo debe contar con una introducción donde el estudiante contextualiza los datos, lo cual le permite profundizar en el área de aplicación donde fue realizado el estudio.
 - Los resultados deberán ser expuestos en forma oral.

6. Evaluación

Se realizarán tres exámenes parciales. Además los estudiantes presentarán un trabajo de análisis de datos y un trabajo adicional.

Si un estudiante faltase a algún examen por **causa justificada**, debe solicitar por escrito la reposición del examen indicando las razones de la ausencia, acompañada de los documentos justificantes. La misma debe entregarse ante el profesor **a más tardar en cinco días hábiles del reintegro a lecciones.**

Exámenes	75 %
Primer examen parcial	20%
Segundo examen parcial	25%
Tercer examen parcial	25%
Trabajos	30 %
Trabajo de análisis de datos	15%
Trabajo adicional	15%





7. Cronograma

	Módulo	L	J	S	Actividad	
AGOSTO		13			Presentación	
	I. Modelo de regresión lineal		16		Lab.0 - DESCRIPTIVO	
			20		Exposición temas 1.1 - 1.3	
			23		Lab. No.1 - ESTIMACION DE COEFICIENTES	
			27		Exposición temas 1.4 - 1.5 (simulación)	
	30		Lab. No.2 - INTERVALOS DE CONFIANZA PARA LOS BETAS			
SEPTIEMBRE		3			NO HABRA LECCIONES	
		6				
		10			Exposición tema 1.6	
			13		Lab. No.3 - INTERVALOS PARA LA MEDIA Y RESPUESTA INDIVIDUAL	
			17		Exposición tema 1.7 - 1.8	
			20		Lab. No.4 - COMPARACION DE MODELOS ANIDADOS	
			22		EXAMEN I: capítulo I (hasta 1.6)	
			24		Exposición tema 1.9	
		27		Lab. No.5 - PREDICTORES CATEGORICOS		
OCTUBRE	II. Selección de predictores	1			Exposición temas 2.1 y 2.2	
			4		Lab. No.6 - SELECCION DE VARIABLES	
	III. Diagnósticos		8			Lab. No.7 - DIAGNOSTICOS I
				11		Exposición temas 3.2 y 3.3
			15			FERIADO
		18		Lab. No.7 - DIAGNOSTICOS I (continuación)		
		22		Lab. No.8 - VALORES EXTREMOS Y CASOS DE INFLUENCIA		
NOVIEMBRE	IV. Construcción del modelo		25		Exposición temas 4.1 - Lab. No.10 - TRANSFORMACIONES	
			29		Exposición temas 4.2 - 4.3	
			1		Lab. No.10 - MINIMOS CUADRADOS PONDERADOS	
			3		EXAMEN II: capítulos II y III + 1.6 A 1.9	
		5		Exposición tema 4.4 - Lab. No.11 - REGRESION ROBUSTA		
		8		Lab. No.12 - COMPONENTES PRINCIPALES		
NOVIEMBRE	V. Modelos lineales generalizados		12		Exposición tema 5.1	
				15		Lab. No.13 - LOGISTICA
			19		Conclusión capítulo V	
			22		Presentación de trabajos	
			26			
			29		EXAMEN III: capítulos IV y V	
DIC			3			
				6		
			10		EXAMEN AMPLIACION	



8. Referencias bibliográficas

Chatterjee, S. & A.S. Hadi (1988). *Sensitivity Análisis in Linear Regression*. New York, N.Y.: Wiley.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536 Ch495s c.1

Faraway, J. (2002). *Practical Regression and Anova using R*. Documento en internet: <http://www.biostat.jhsph.edu/~iruczins/teaching/jf/faraway.html>

Faraway, J. (2004). *Linear Models with R*. Chapman & Hall/CRC.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536 F219L

Faraway, J. (2005). *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Chapman & Hall/CRC.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536 F219e

Fox, J. (2011). *An R ans S-Plus Companion to Applied Regression*. 2a edición. Thousand Oaks, Cal.: Sage Publications, Inc.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536.028.551.33 F792a2

Gujarati, D. (1990). *Econometría básica*. McGraw-Hill.

BIBLIOTECA LUIS DEMETRIO TINOCO 330.015.195 G969e2

Kutner, M, Nachtsheim, C, Neter, J, Li, William (2005). *Applied Linear Statistical Models*. 5a. ed. WCB, McGraw-Hill.

BIBLIOTECA LUIS DEMETRIO TINOCO 519.536 K97a5